CSE 840: Computational Foundations of Artificial Intelligence Nov. 13, 2023

Lebesgue decomposition, CDF, Random variables

Instructor: Vishnu Boddeti *Scribe: Jingzhe Liu, Hongzhi Wen, Yufeng Li*

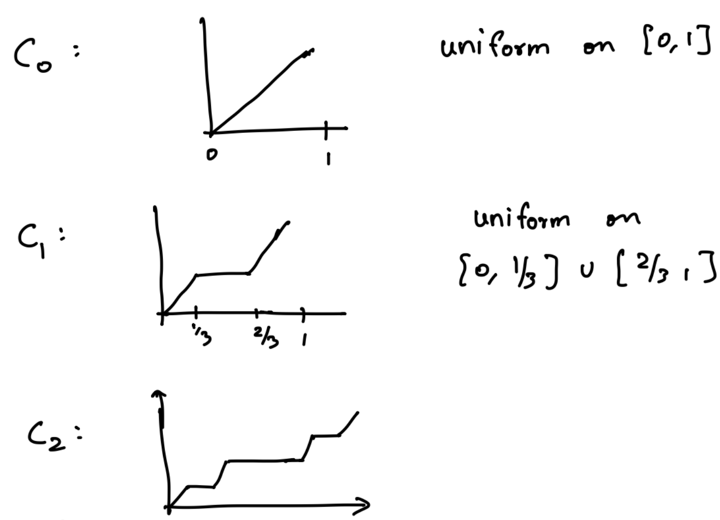
1 Lebegue Decomposition

Definition 1 *Consider a measure space $(X, \mathcal{A}, \lambda)$ and anthor measure $\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$.*
(a) μ is called absolutely continuous if $\lambda(A) = 0 \Rightarrow \mu(A) = 0$ for all $A \in \mathcal{B}((R)$.
(b) μ is called singular with respect to λ if there is $N \in \mathcal{B}(\mathbb{R})$ with $\lambda(N) = 0$ and $\mu(N^c) = 0$

Theorem 2 *Consider μ, γ prob. measures on $(\Omega, \mathcal{A})$. Then there exists a unique decomposition $\gamma = \gamma_{ac} + \gamma_s$ such that $\gamma_{ac} \ll \mu$ and $\gamma_s \perp \mu$.*

Example: $\gamma = \frac{1}{2}(N(0,1), \delta_0)$. $\gamma = \gamma_{ac} + \gamma_s$, where $\gamma_{ac} = \frac{1}{2}N(0,1)$, $\gamma_s = \frac{1}{2}\delta_0$.
Cantor distribution: non-trivial distribution that is singular with respect to λ. Construct the Cantor set:
· Start with $C_0 := [0, 1]$
· Start with $C_1 := [0, \frac{1}{3}] \bigcup [\frac{2}{3}, 1]$
· · ·

The Cantor set is limited in this process. Now construct a prob. distribution:
Consider the CDFs of the sets C_0, C_1, C_2 · · ·

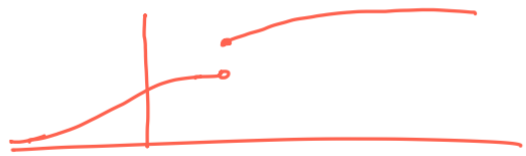


⋮

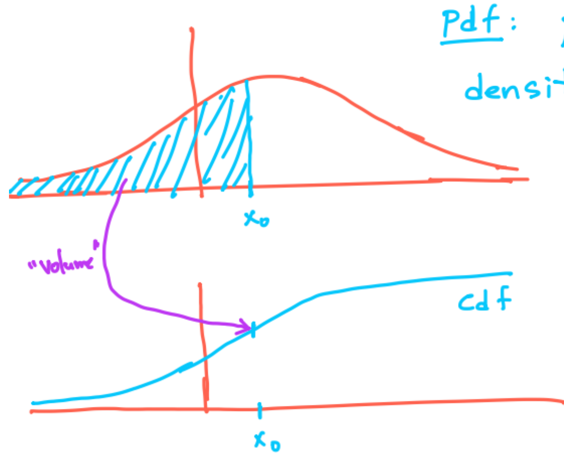
Take limit $\to T$. Can prove many interesting properties.

2 Cumulative Distribution Function

Let P be a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Define the function $F : \mathbb{R} \to \mathbb{R}$, $x \to P((-\infty, x])$. W says that F is a cumulative distribution function(cdf), that satisfies the following properties:
(i) F is monotonically increasing, $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to \infty} F(x) = 1$
(ii) F is continuous from the right: $(x_n)_{n \in \mathbb{N}}$ sequence with $x_n \leq x_{n+1}$ and $x_n \to x$ then also $F(x_n) \to F(x)$



Pdf: prob. density func.
density of normal dist.



Let $F : \mathbb{R} \to \mathbb{R}$ be a function with properties (i) and (ii). Then there exists a unique probability measure P on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $P((-\infty, x]) := F(x)$

3 Random Varibale

Definition 3 *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(\widetilde{\Omega}, \widetilde{\mathcal{A}})$ be another measuable space. A mapping $X : \Omega \to \widetilde{\Omega}$ is called a random varibale if X is measurable, i.e., $\forall \widetilde{A} \in \widetilde{\mathcal{A}} : X^{-1}(\widetilde{A}) := \{w \in \Omega | X(w) \in \widetilde{A}\} \in \mathcal{A}$.*

Definition 4 *A random variable $X : \Omega \to \widetilde{\Omega}$ induces a measure on the target space: For $\widetilde{A} \in \widetilde{\mathcal{A}}$ we define $P_X(\widetilde{\mathcal{A}}) := P(X^{-1}(\widetilde{A}))$. This is a probability measure on $(\widetilde{\Omega}, \widetilde{\mathcal{A}})$, and it is called the distribution of X.*

Definition 5 $X : (\Omega, \mathcal{A}, P) \to (\widetilde{\Omega}, \widetilde{\mathcal{A}})$. *Then the family $\sigma(X) := \{X^{-1}(\widetilde{A}) | \widetilde{A} \in \widetilde{\mathcal{A}}\}$ is a σ-algebra induced by X. (It is the smallest σ-algebra on Ω that makes X measurable)*

4 Conditional Probability

Notation: $P(A \bigcap B) = P(\text{"A and B"})$; $P(A \bigcup B) = P(\text{"A or B"})$

Definition 6 *Let $(\Omega, \mathcal{A}, P)$ be a probability space. $A, B \in \mathcal{A}$, $P(B) > 0$. Then $P(A|B) := \frac{P(A \bigcap B)}{P(B)}$ is called the conditional probability of A given B.*

Theorem 7 *The mapping $P_B : \mathcal{A} \to [0, 1]$, $A \to P(A|B)$ is a probability measure on $(\Omega, \mathcal{A})$, it is called the conditional distribution of P with respect to B.*

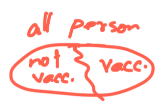
Examples:

1. two dice: $P(\text{"sum is q"}|\text{"first die was 3"})$
2. $\Omega =$ all persons on earth, $\mathcal{A} = P(\Omega)$, $P =$"uniform".

event A: "person has been vaccinated

B: "person has disease"

P(disease | vaccinated)



P (vaccinated | disease)